

Demonstration-based training (DBT) in the design of a video tutorial for software training

Hans van der Meij¹  · Jan van der Meij²

Received: 14 April 2016 / Accepted: 1 September 2016 / Published online: 8 September 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract This study investigates the design and effectiveness of a video tutorial for software training. In accordance with demonstration-based training, the tutorial consisted of a series of task demonstrations, with instructional features added to enhance learning. An experiment is reported in which a demonstration-only control condition was compared with a demonstration-with-review condition. The review videos provided a recap to support retention of the procedure. Strict viewing conditions were imposed during training. Both learning and motivation were assessed. Participants were 55 students from primary and secondary schools in Germany. The outcomes showed that both tutorials significantly contributed to development of procedural knowledge and to motivation, and that the review tutorial surpassed the control. The discussion addresses the role of user control. In addition, with the effectiveness of reviews apparently becoming better established, the need arises for measures that can reveal the effects of reviews on observational learning processes.

Keywords Demonstration-based training (DBT) · Software training · Tutorials · Instructional video · Procedural knowledge development · Reviews

✉ Hans van der Meij
h.vandermeij@utwente.nl

Jan van der Meij
j.vandermeij@utwente.nl

¹ Department of Instructional Technology, Faculty of Behavioural, Management and Social Sciences, University of Twente, Drienerlolaan 5, 7522 AE Enschede, The Netherlands

² Faculty of Behavioural, Management and Social Sciences, ELAN, Institute for Teacher Education and Science Communication, University of Twente, Drienerlolaan 5, 7522 AE Enschede, The Netherlands

Introduction

Recent years have witnessed a rapid increase in the popularity of instructional video (Giannakos 2013). One contributor to this expanded usage is that software companies have begun to produce and distribute video rather than paper tutorials. The goal of a software tutorial is for the user to develop procedural knowledge. Procedures are lists of steps or algorithms (Mayer 2008). The user must get to know the sequence of activities that lead to task completion in a particular software program. This requires the user to learn to perform a series of actions that lead to observable changes in the software program.

An important way for people to acquire knowledge about task completion is by observing a model of performance. The demonstration alone is insufficient for learning how to complete a task, however. Instructional support is necessary in order to turn a demonstration into an effective means of learning. One framework used to accomplish this is demonstration-based training (DBT). DBT is a design approach in which a dynamic example of performance is complemented with instructional features (Grossman et al. 2013; Rosen et al. 2010).

The basis for DBT is Bandura's (1986) theory of observational learning. This theory posits that observational or model-based learning hinges on the four interrelated processes of attention, retention, production, and motivation. The design principles advanced in DBT are explicitly or implicitly coupled with these processes. The tutorial in the present study was developed in accordance with DBT. To our knowledge, a DBT-based approach has rarely been adopted for the construction of a video tutorial for software training. This research should therefore be taken as an exploration of design options and their effectiveness for software training with video.

The following section describes the four fundamental processes in observational learning, along with design guidelines that can enhance these processes in a video tutorial for software training. Special attention is given to the issue of following a task demonstration with a review in order to enhance learning from the tutorial. An experiment is reported in which the effectiveness of a demonstration-based video tutorial with or without reviews is investigated.

Demonstration-based-training (DBT)

Observational learning hinges on the interrelated processes of attention, retention, production, and motivation (Bandura 1986).

Attention Attention is a complex mechanism that determines what information gets past an observer's initial information processing stage (Anderson 2010). It is the process in which some information is enhanced while other information is inhibited. Attention can be a top-down or bottom-up process. Attention is a top-down process when prior knowledge influences the information that a user perceives and to which a user attends. It is a bottom-up process when physical features of the video draw the users' attention.

During a demonstration the designer has no influence over top-down attentional processes because these depend on the user's prior knowledge. Instructional measures included in a demonstration therefore always concern bottom-up attentional processes. However, these measures may not always suffice, in which case pre-training can be helpful. In pre-training users are acquainted with the names and characteristics of the main

concepts involved in a demonstration, equipping the user with knowledge that makes it easier to process the demonstration.

That even a relatively short period of pre-training can affect attentional processes in a top-down fashion was shown in the research of Canham and Hegarty (2010). In the study, psychology students were first given the complex task of judging the wind direction from reading a weather map after which they received a 10–15 min instruction in meteorology. Thereafter, they were given new judgment tasks for predicting the wind direction from weather maps. The data showed that the instructions significantly affected the participants' attentional processes. After the instruction, the participants spend more time viewing the task-relevant information on the maps and less time on task-irrelevant information. Encoding too was affected and also there was improved success on a transfer task. More generally, a considerable number of empirical studies have found that pre-training enhances learning from multimedia presentations (Mayer and Pilegard 2014).

An instructional measure that can support bottom-up attention processes is cueing or *signaling*. Signaling is a design-based technique for drawing attention to key points of information (Lemarié et al. 2008). Signals do not add content. Rather, they emphasize structure, locations, or objects by making them stand out. Signals support attentional processes in a bottom-up fashion.

A multimedia presentation can provide verbal and/or visual cues. Examples of verbal signals are deictic references (e.g., “On the right hand side ...”), intonation, and key words that stress relevance (e.g., “Pay special attention”, and “This is important”). To our knowledge little research has assessed the attentional effect of such verbal cues. In contrast, there is considerable research on visuals that can draw attention. Features that can affect the selection of attention are color, sudden appearance and movement (e.g., de Koning et al. 2009; Jamet 2014; Kosslyn et al. 2012). More generally, large perceptible differences draw attention (e.g., drawing a red circle around a key object from an interface).

Empirical evidence of the attention-supportive role of signaling comes from eye-tracking data in research on animations. Several studies have revealed that users fixate more often and spend more time on cued information (e.g., Boucheix and Lowe 2010; de Koning et al. 2010; Kriz and Hegarty 2007). Equivocal findings have been reported for the contribution of signaling to learning outcomes. Some studies have found significant effects (e.g., Amadiou et al. 2011; Boucheix and Guignard 2005; Jin 2013), while others have found no significant effects on learning outcomes even when the signals effectively served their attention-guiding role (e.g., de Koning et al. 2010; Kriz and Hegarty 2007; Skuballa et al. 2012). None of these studies reported negative effects of signaling on learning, however.

Retention Retention involves the process of transforming incoming information into symbolic codes that are stored in long-term memory (Bandura 1986). The observer must extract the distinctive features and structures from the modeled activities. The result of these efforts should be a succinct, prototypical representation of task performance. The coded information must be further committed to memory to serve as a guide for subsequent action. Retention thus includes key processes of information processing that Mayer (2008) refers to as organization and integration. The end result of the retention process should be a concept of task performance that can serve as a guide and standard for future actions, enabling the user to organize, initiate, and monitor these actions (Bandura 1986).

An instructional measure that can support the organizational aspect of retention is *segmentation*. Segmentation involves dividing a continuous animation into smaller units or sections with a beginning and end. Event theory provides an important framework for

constructing such segments. According to event theory (Zacks and Tversky 2001, 2003), people tend to conceive of activities as discrete events, which thus determine understanding and memory. This process can be capitalized on and enhanced by structuring a procedure into meaningful segments.

Ertelt (2007) found empirical support for the contribution of segmentation to learning from a video tutorial for software training. The segmentation was content-based. More specifically, every solution step was given a label, effectively creating a series of sub-goals for a complex procedure. Labelling was found to contribute significantly to participants' knowledge development. Positive effects of segmentation on learning have likewise been reported in other empirical studies (e.g., Boltz 1992; Catrambone 1995, 1998; Margulieux et al. 2012; Schitteck Janda et al. 2005; Schwan et al. 2000).

An instructional measure that can support the integrative aspect of retention is the inclusion of *pauses*. The transient nature of an animation can be taxing, making it difficult to connect incoming information with what is already known (Brucker et al. 2014; Lowe et al. 2011). A design measure that can moderate this problem is the inclusion of brief pauses of 2–5 s. Pauses are time-stamped demarcations that disrupt the continuous information stream. They give the user a brief respite to digest what has been observed, in which new information can be connected with known information (Spanjers et al. 2010).

Two experiments by Moreno (2007) found that pauses significantly affected participants' appraisals of the processing effort required by learning materials (i.e., video and animation). Participants judged animations with pauses as less difficult and requiring less mental effort than animations without pauses. In addition, the paused materials led to significantly higher scores on a transfer test. Similar findings favoring pauses have been reported in other empirical studies (e.g., Hasler et al. 2007; Hassanabadi et al. 2011; Spanjers et al. 2012; Spanjers et al. 2011).

Production The ultimate goal of observing a demonstration is for the user to be able to accomplish a task that is similar or related to the one demonstrated. Production refers to the user's capacity to execute the steps in a procedure correctly so that task completion is accomplished (Bandura 1986).

An instructional measure that can support production is *practice*. Empirical support for complementing animations with practice to enhance learning has been found in several studies (e.g., Ertelt 2007; Leppink et al. 2014; Rieber 1990; van Gog et al. 2011). Two explanations for the effectiveness of practice are generally offered. One argument is that practice stimulates reflection. Observational learning runs the same risk of passive and superficial processing noted for worked examples (Atkinson et al. 2000). Practice can counter this effect by stimulating deeper processing (Ertelt 2007). After having seen a model of performance, practice can serve as a check of understanding. The other argument is that practice consolidates learning. The modeled performance provides users with a mental model of the solution process which can then subsequently be strengthened by practicing with a similar problem (Hodges and Coppola 2014; van Gog 2011).

An alternative explanation for the contribution of practice to learning can be found in research on the testing effect (e.g., Karpicke and Roediger 2008; Roediger and Karpicke 2006). These studies have challenged the prevalent view that retrieving information on a test is considered a relatively neutral event and that learning only occurs when people study and encode material. The standard assumption is that testing merely measures the learning that has occurred and does not by itself produce learning. A considerable number of studies on the testing effect have revealed otherwise. They have found that testing can yield better retention in the long run than restudying (Karpicke and Roediger 2007; Larsen et al. 2013). One explanation offered for the testing effect is that it consolidates learning by

strengthening the memory trace (e.g., Brewer et al. 2010). Recent findings of a repeated testing effect on transfer items further suggest that there can also be a contribution to deep learning (e.g., Butler 2010; Dirx et al. 2014; McDaniel et al. 2013).

A positive effect of practice is not universally found, however. Some empirical studies have pointed to the moderating role of prior knowledge (Reisslein et al. 2006; Wouters et al. 2010). In the study by Reisslein et al., the factor of prior knowledge was even systematically varied to ascertain its influence, revealing that participants with low prior knowledge benefitted most from the classic instructional set-up in which instructions were followed by practice. In contrast, high prior knowledge participants performed better on near (but not far) transfer test items when practice preceded the instructions. In other words, the study revealed that practice benefits low and high prior knowledge participants when it is appropriate for their stage of learning.

Motivation The driving force behind the processes of attention, retention, and production is motivation. It can be defined as the process whereby goal-directed activity is instigated and sustained (Pintrich and Schunk 2002). An attractive framework for constructing instructional features that can increase motivation in performance-related situations is the CANE model (Clark 2015). According to this model, motivation stems from the interlinked processes of commitment and mental effort. The factors that stimulate people to commit themselves to a task, and engage in its active and sustained pursuit, are self-efficacy, mood, and personal goal value. The factors that propel people to spend effort on task completion have to do with effectiveness values, which include utility, interest, and importance (Eccles and Wigfield 2002).

An instructional measure that can enhance task commitment is a *simple-to-complex* task sequence. When task complexity increases gradually instead of sharply, each new task confronts the user with a new and attainable challenge. The principle of adopting a simple-to-complex sequence is generally considered to be an effective design measure for optimizing the users' cognitive load during problem solving (e.g., Kester et al. 2001; van Merriënboer, 1997). That such a sequence can also help increase the user's motivation is less often noted. More specifically, we would like to argue that a simple-to-complex sequence supports the user's development of self-efficacy. Self-efficacy is the users' expectancy for success on a novel task (Bandura, 1997). With a simple-to-complex task sequence, the user is optimally prepared for dealing with the challenge posed by new task demands.

An instructional measure that can enhance the expenditure of mental effort is the *anchoring of the tools in the task domain*. One of the key principles of the minimalist approach to software training is that users should be given tasks that can instantly be recognized as genuine and valuable (van der Meij and Carroll 1998). Users are not motivated by tasks that merely teach them how to use the tools provided by the software, but they can probably be motivated by tasks that, to the greatest extent possible, are core tasks for which they might want to use the software in daily life.

Experimental design and research questions

Learning from a video tutorial can probably be enhanced by the added presence of a review. A review is a short recap of the main steps in a procedure. It reiterates what is involved in task completion in an abbreviated form, and as such it should support retention.

As the name implies, a review is presented after a demonstration. Therefore, its presence does not affect the design of the demonstration.

Reviews have received little attention in multimedia research. A recent experimental study (van der Meij and van der Meij 2016) showed that a video tutorial for software training that followed up task demonstrations with a review yielded significantly higher learning outcomes on an immediate post-test than did the same tutorial without reviews. However, no effects were found on the training tasks, or on a delayed post-test. Both tutorials were found to substantially increase motivation (i.e., task relevance and self-efficacy). In addition, the demonstration-with-review tutorial increased participants' self-efficacy significantly more than did the demonstration-only tutorial.

The present study further investigates the effects of a review on learning from a video tutorial. It does so by imposing strict control over video play. In the aforementioned study, participants had user control. During training they could use a toolbar to play, pause, and replay the videos as they deemed fit. Active use of such control during initial intake and during practice may have obscured the effectiveness of the reviews. Therefore, the present study put control of video play in the hands of the experimenter.

The study is quasi-experimental. The control condition presents a video tutorial that consists only of task demonstrations that have been enhanced with all of the aforementioned instructional measures. The experimental condition presents the same tutorial plus reviews. The experiment includes a pre-test and it assesses procedural knowledge development during and after training. All these learning measures are performance tests involving the same type of tasks as demonstrated. In addition, participants' motivation (i.e., task relevance and self-efficacy) is measured. The following research questions are addressed:

- Question 1 How well do the tutorials support learning?
- Question 2 Has the review tutorial a stronger effect on learning than the control tutorial?
- Question 3 How effective are the tutorials in raising motivation?
- Question 4 Has the review tutorial a stronger effect on motivation than the control tutorial?

Based on earlier studies (van der Meij 2014; van der Meij and van der Meij 2014, 2015, 2016) the tutorials were expected to contribute significantly to procedural knowledge development when comparing scores for practice, immediate and delayed post-test to pre-test scores. In addition, we expected participants in the experimental condition to demonstrate better performances on these measures than participants in the control condition.

Based on earlier research (van der Meij 2014; van der Meij and van der Meij 2014, 2015, 2016) the tutorials were expected to yield higher scores for appraisals of task relevance and self-efficacy after training in comparison to motivation measures taken before training. No effect of condition on task-relevance was expected because the reviews did not convey new information on task relevance. A significant effect of condition on self-efficacy development was expected.

Method

Participants

There were 55 participants in the study. Ten participants came from the highest grade-level classroom of an elementary school in Germany. Forty-five participants came from the first

and second grade classrooms of a secondary school in Germany. The mean age of the 24 male and 31 female participants was 11.4 years (range 8.7–13.9). Participants were randomly assigned to the control or experimental condition, after stratification for school and classroom. All instructional materials, including the software, were in German.

Instructional materials

The video tutorials taught formatting tasks in Microsoft Word (2007 version). The tasks were anchored in the audience's task domain. That is, the videos presented examples of (sections of) school reports that benefitted from having the demonstrated formatting tasks completed. A table of contents with titles and subtitles organized the tasks (see Fig. 1). This structure created a meaningful segmentation of the various formatting tasks. Based on the literature providing advice on videos (e.g., Guo et al. 2014; Plaisant and Shneiderman 2005; Wistia 2012) an upper limit of 3 min was chosen for individual videos. The order of task presentation followed a simple-to-complex sequence. The first chapter discussed adjusting the left and right margins for an entire document. The second chapter revolved around formatting paragraphs, citations, and lists. The last chapter dealt with the automatic creation of a table of contents.

Demonstration videos were recorded demonstrations. They displayed animated sequences of screenshots with narration. Each demonstration began with a goal statement. For instance, the video on adjusting the right margin began as follows: "The margins in this text are too small. There is not enough space between the words and the page border. We will start by adjusting the right margin." Thereafter, the demonstration showed and told how to complete the particular formatting task. Demonstrations occasionally included short 2-s pauses between sections, and always ended with such a pause. Visual signals regularly indicated pertinent interface locations and objects. The mean length of a task demonstration video was 1.14 min (range 0.48–1.46).

Review videos summarized task completion. Each task review followed automatically after the demonstration for that task, beginning after the 2-s pause at the end of the demonstration. Reviews were indicated with the following introductory statement: "You are now done, but remember ..." Reviews lasted between 13 and 26 s. The mean length of a video with review was 1.31 min (range 1.02–2.06).

Instruments

Tests Four tests (i.e., pre-test, practice, immediate post-test, delayed post-test) assessed learning. The test items presented formatting tasks similar to those in the task demonstration videos. A score of 0 points was awarded for each item the participant could not complete correctly. Correct task completion yielded a score of 1. Scores were converted to a percentage of possible points. Reliability analyses, Cronbach's alpha, led to moderate results for all knowledge tests: pre-test ($\alpha = 0.53$), training ($\alpha = 0.70$), immediate post-test ($\alpha = 0.57$), and delayed post-test ($\alpha = 0.54$).

Practice files During training, the participants were given an opportunity for practice after watching each task demonstration. Because they had no access to the video, their practice was also a test of performance. To facilitate task effort during training, special practice files were constructed that minimized the need for task-irrelevant actions, such as typing text. In addition, these files included a minimum of distracting features (see van der Meij and Carroll 1998). Practice files also standardized practice; they made task

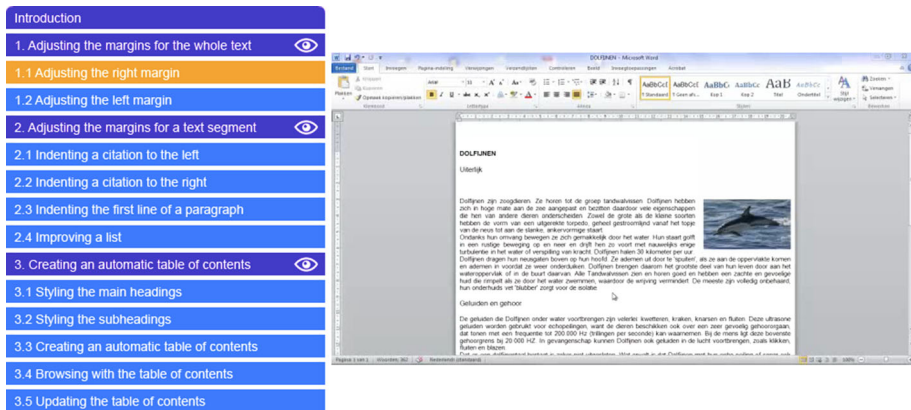


Fig. 1 Screenshot of the website (translated version)

completion efforts comparable across conditions. Practice files were accessible from a folder with the student's name that was on the computer desktop.

Initial experience and motivation questionnaire (IEMQ) The IEMQ was a paper-and-pencil instrument that presented a screenshot and asked a set of three questions for each of the six training tasks: (a) "Have you ever seen this?" (Experience), (b) "How often do you need to complete this task?" (Relevance), and (c) "How well do you think you can complete this task?" (Self-efficacy). Answers were given on a 7-point Likert scale ranging from never (1) to always (7), or very poorly (1) to very well (7). Reliability analyses using Cronbach's alpha indicated satisfactory results (i.e., Experience $\alpha = 0.81$; Relevance $\alpha = 0.88$; Self-efficacy $\alpha = 0.82$).

Final motivation questionnaire (FMQ) The FMQ was a paper-and-pencil instrument. There were eight questions about task relevance (e.g., "I can use what I learned in various ways for schoolwork", and "I think it is important to present lists in a well-structured manner"), and eight questions about self-efficacy (e.g., "I can now present a nicely structured list", and "I now know how to indent the first line of a new text segment"). Answers were given on a 7-point Likert scale. Reliability scores using Cronbach's alpha were 0.92 for task relevance and 0.81 for self-efficacy.

Both motivation questionnaires employed in the study had their origin in the Fragenbogen Aktuelle Motivation [Questionnaire on Current Motivation.] (Rheinberg et al. 2001). They were adapted for software training and were used in earlier empirical studies (van der Meij 2014; van der Meij and van der Meij 2014, 2015, 2016).

Procedure

The experiment was conducted in three sessions that took place in the schools' computer rooms. One room was reserved for the participants from the experimental group, while the participants from the control group were instructed in the other room. The first session began with a 5-min introduction, after which the IEMQ and pre-test were completed (20 min maximum). Training took place 1 or 2 days later, with separate sessions for the control and experimental groups. Each training session began with a whole group introduction (10 min). All videos used in the training were presented to the group as a whole via a screen in front of the room. After each task demonstration (with review), participants had

5 min to complete a practice test item. After training was finished there was a 5-min break, followed by the FMQ (3 min) and the immediate post-test (20 min). Two weeks later, participants took the delayed post-test (20 min).

Analysis

Repeated measures ANOVAs were conducted to examine the effects of time and condition on test scores and motivation. Tests were one-tailed for directional predictions, and two-tailed for other cases, with alpha set at 0.05. The degrees of freedom occasionally differed due to missing data. Cohen's (1988) d -statistic is used to report effect size. These tend to be qualified as small for $d = 0.2$, medium for $d = 0.5$ and large for $d = 0.8$.

Results

Learning

A repeated measures ANOVA with *test scores* as the dependent variable indicated that there was no interaction between *time* \times *condition*, $F(3, 159) = 1.10$, *ns*.

There was a significant main effect of *time*, $F(3, 159) = 58.23$, $p < 0.001$. Table 1 shows that, compared with pre-test scores, participants' scores after working with the tutorials were higher at all three subsequent testing points. Even the lowest of these three scores (i.e., delayed post-test) was significantly higher than the pre-test score, $F(1, 53) = 41.72$, $p < 0.001$, $d = 1.15$. Pairwise comparisons showed that there was a significant increase from pre-test to practice ($F(1, 53) = 173.35$, $p < 0.001$, $d = 2.29$), that there was a significant decrease from practice to immediate post-test ($F(1, 53) = 38.37$, $p < 0.001$, $d = 1.02$), and that scores remained the same for the two post-tests ($F < 1$).

There was a significant main effect for *condition*, $F(1, 53) = 11.47$, $p = 0.001$, $d = 0.92$. Separate ANOVAs for the four measurement time points revealed no difference on the pre-test, $F(1, 54) = 1.58$, *ns*. Condition had a significant effect on the practice test score, $F(1, 54) = 9.55$, $p = 0.002$, $d = 0.83$, on the immediate post-test score, $F(1, 54) = 3.90$, $p = 0.027$, $d = 0.53$, on the delayed post-test score, $F(1, 54) = 3.63$, $p = 0.031$, $d = 0.52$. All of these comparisons favored the review tutorial (see Table 1). The slight difference between conditions at the start prompted us to also conduct ANCOVAs with the pre-test scores as the covariate. These confirmed the ANCOVA findings. That is, the experimental condition outperformed the control on the practice test ($F(1, 52) = 8.47$, $p = 0.002$), the immediate post-test, ($F(1, 52) = 3.37$, $p = 0.036$), and the delayed post-test, ($F(1, 52) = 3.01$, $p = 0.045$).

Table 1 Mean success rates for tests by condition

Condition	Pre-test M (SD)	Training M (SD)	Immediate post-test M (SD)	Delayed post-test M (SD)
Control (n = 27)	15.3 (16.4)	58.3 (24.3)	36.6 (22.4)	36.1 (21.5)
Review (n = 28)	21.4 (18.0)	78.1 (23.2)	49.1 (24.5)	47.8 (23.8)
Total (n = 55)	18.4 (17.3)	68.4 (25.6)	43.0 (24.1)	42.1 (23.2)

Motivation

A repeated measures ANOVA with *task relevance* as the dependent variable showed that there was no interaction between *time* \times *condition*, $F(1, 47) = 2.86$, $p = 0.097$.

There was a significant main effect of *time*, $F(1, 47) = 106.40$, $p < 0.001$, $d = 1.74$. Table 2 shows that the tutorials yielded higher appraisals for task relevance after training than before.

There was also a significant main effect for *condition*, $F(1, 47) = 5.66$, $p = 0.021$, $d = 0.90$. Scores for task appraisals before training did not differ by condition, $F(1, 53) = 1.32$, *ns*. However, after training, participants viewing the review tutorial gave significantly higher appraisals, $F(1, 49) = 9.96$, $p = 0.003$, $d = 1.70$.

A repeated measures ANOVA with *self-efficacy* as the dependent variable showed that there was no interaction between *time* \times *condition*, $F(1, 47) < 1$.

There was a significant main effect of *time*, $F(1, 47) = 49.61$, $p < 0.001$, $d = 1.25$. Table 2 shows that the tutorials yielded higher appraisals for self-efficacy after training than before.

There was no main effect for *condition*, $F(1, 47) = 2.98$, $p = 0.091$. Scores for self-efficacy did not differ before training, $F(1, 53) = 2.77$, *ns*, or after training, $F(1, 49) < 1$.

Discussion

Comparisons between pre-test scores on the one hand, and practice and post-test scores on the other, revealed that the tutorials yielded significant and substantial learning. In short, the findings show that the tutorials effectively supported procedural knowledge development. This outcome further substantiates earlier research that also reported significant learning gains in software training with DBT-based video tutorials (van der Meij 2014; van der Meij and van der Meij 2014, 2015, 2016).

An explanation for this outcome derives from the coupling of instructional measures to fundamental processes in observational learning. Specifically, the video tutorials presented users with a model of performance for task accomplishment to which design features were added to enhance learning. Some of these measures concerned the task demonstrations themselves, such as the inclusion of verbal and visual cues for drawing user attention.

Table 2 Mean scores for task relevance and self-efficacy by condition

Condition	Task relevance		Self-efficacy	
	Before M (SD)	After M (SD)	Before M (SD)	After M (SD)
Control ($n_{\text{Before}} = 27$ $n_{\text{After}} = 24$)	2.19 (1.23)	4.31 (1.29)	2.59 (1.41)	4.49 (1.18)
Review ($n_{\text{Before}} = 27$ $n_{\text{After}} = 25$)	2.66 (1.74)	5.50 (1.35)	3.29 (1.66)	4.82 (1.25)
Total ($n_{\text{Before}} = 54$ $n_{\text{After}} = 49$)	2.42 (1.51)	4.93 (1.44)	2.94 (1.57)	4.66 (1.22)

Scale maximum is 7. A higher score indicates a more positive evaluation

Other measures were complementary such as the inclusion of practice on a task immediately after a video demonstration.

The outcomes were not uniformly positive, however. The significant decline from practice to post-tests showed that a considerable part of what was learned was also quickly forgotten. The posttest scores for both tutorials were also lower than what was found in earlier studies with similar DBT-based video tutorials for software training (van der Meij 2014; van der Meij and van der Meij 2014, 2015, 2016). What can account for this finding? The most likely explanation rests with the restricted viewing conditions.

The restriction was imposed to better assess the contribution to learning of the reviews. The first study that investigated the contribution of reviews in video tutorials found a positive effect on an immediate but not on a delayed post-test (van der Meij and van der Meij 2016). The limited effectiveness of the review was attributed to user control playing an intermediary role. That is, it was argued that participants could stop and replay videos during training as they deemed fit, and therefore needed to rely less on the reviews. The present study therefore restricted the viewing conditions.

Video play was strictly experimenter-controlled; all videos were shown to the participants only once, in a system-paced mode. In addition, there was no video access during practice. Immediately after a task video was shown, the participants were asked to complete a similar task on their own, without the opportunity to look at the video again. So, how did the restricted viewing reduce recall over time?

One explanation is that the lack of user control affected information intake during video viewing, which subsequently influenced learning. The combination of a complex user interface and an ongoing stream of text and images can be taxing. Indeed, it is possible that the presentations occasionally placed too high a demand on the users' working memory. Working memory is limited in the information it can retain and process (Baddeley, 2007), which should prompt designers to include instructional measures that can reduce processing load.

Kosslyn et al. (2012) proposed the principles of limited capacity and informative change, and suggested corollary design measures for each. The first principle states that users are restricted in how much information they can retain and process. Design measures such as signals and pauses can address this limitation. The second principle holds that people expect changes of information to be conveyed by perceptual properties of the learning material. A design measure that can help users cope with such changes is segmentation. With segmentation the beginning and end point of sections are clearly demarcated for the user.

In addition, the designer can equip the video tutorial with a toolbar that gives viewers the capability of pausing the video and replaying sections when needed. With such a console, the viewers themselves can take action when there is a risk of cognitive overload. Empirical research shows that users sometimes spontaneously use the toolbar so that their capacity for information intake and video viewing are well-aligned. For instance, Schwan and Riempp (2004) reported heavier use of the interactive features of the toolbar when users were viewing more difficult tasks. Merkt et al. (2011) reported similar adaptive behaviors by the user.

Several empirical studies have further shown that user control can have a positive effect on learning (e.g., Hasler et al., 2007; Höffler and Schwartz 2011; Schwan and Riempp 2004; Stiller et al. 2009). Such effects are not consistently found, however. There are also a considerable number of empirical studies that failed to find a positive effect of user control on learning (e.g., Boucheix and Guignard 2005; Chen, 2014; Kriz and Hegarty 2007; Lowe, 2004; Pedra et al. 2015; Tabbers and De Koeijer 2010). These studies point to the

influence of prior knowledge as a mediating factor, and they speak to the functional usage of the control options.

Another explanation is that user control directly affects what takes place during practice, which, in turn, affects learning. An empirical study by Shippey et al. (2011) is illustrative. The study compared independent practice, supervised practice, and practice with video access. The audience consisted of medical students who needed to learn a surgical procedure. The outcomes revealed that sustained performance results were obtained only in the video condition. The explanation was that the video helped resolve uncertainties about how to execute components of the task. In addition, the video served as a visual reference that enabled the students to assess and adapt their performance.

The present study found strong evidence of the benefits of reviews. The review increased the presentation time for the videos by about 20 %. The positive findings for the review may therefore be due in part to longer exposure to the instructional materials. While this does not necessarily invalidate the claim that the review is beneficial, this is an issue that merits further research.

We believe that the main cognitive process affected by the review is retention. When participants view a demonstration video, they should try to understand what is happening and memorize the procedure. That is, participants must transform what they have seen demonstrated into symbolic codes that are stored for future behavior (Bandura 1986). Because it is impossible to remember everything, the end result is an abbreviated version of the procedure. Reviews can contribute to this process by presenting the procedure in condensed format which then serves as a mental model for later task completion.

Based on this reasoning, a pure repetition of (part of) a demonstration cannot achieve the same effect as a review. However, only experimental testing can tell. In conducting such a test, it seems useful to glean insights from fundamental research on memory representations to better understand how the depth of processing at intake relates to the completeness of the representation on which later task performance is based. The distinction between verbatim and gist memory seems especially relevant (e.g., Abadie et al. 2013; Reyna 2012; Reyna and Brainerd 1995). That is, the repetition of a procedure might have a strong effect on verbatim memory development, which refers to a detailed representation of the exact information. In contrast, the review of a procedure would appear to primarily support the development of gist memory, which refers to the meaning of the information.

The tutorials also had a positive effect on motivation. The findings showed that there was a significant and substantial gain in both task relevance and self-efficacy. This finding is in accord with the results of earlier studies with DBT-based tutorials (van der Meij 2014; van der Meij and van der Meij 2014, 2015, 2016). The effect is ascribed to the design measures taken to support motivation. The simple-to-complex sequencing of tasks is believed to have contributed to self-efficacy. Presenting tasks that are challenging but manageable should support the participants' belief in their capacities in handling formatting tasks in the future (e.g., Clark 2015; Keller 2010). The relevance of the formatting options was conveyed by domain-anchoring, an important design principle in the minimalist approach to software documentation (van der Meij and Carroll 1998). The demonstration videos showed how the successful completion of formatting tasks improved the presentation qualities of a school report. This should enhance perceptions of task relevance (Keller 2010).

The tutorial with reviews also yielded a significantly higher appraisal for task relevance than the control tutorial. This finding was unexpected because the reviews did not present new information that could have piqued user interest in the formatting tasks. Apparently, the concise goal summary provided by the reviews helped increase task relevance. The

predicted effect of condition on self-efficacy was not found, however. An important facet of self-efficacy is the feeling of personal control over the outcome of one's actions (Bandura 1997). In the present study, the viewing conditions were strictly controlled for experimental purposes. Perhaps this lack of user control prevented an effect of reviews on self-efficacy.

All in all, the present study shows that the DBT-based video tutorials substantially contributed to learning and motivation, and that the additional presence of a review further added to these effects. The outcomes warrant reasonable optimism regarding the effectiveness of DBT-based video tutorials with reviews for software training.

A limitation of the present study is the absence of process data. Future research on the effectiveness of video tutorials for software training will need to get a better view of how the design features affect learning. For reviews, we have thus far assumed that their primary impact is on the retention process. Having established some proof that reviews have a positive effect learning, it is time to come to a deeper understanding of the possible causes. To frame it in Schoenfeld's words, "and then you have to study the hell out of it" (1999, p. 12). In other words, future studies of reviews in video tutorials should gather extensive process data and try to obtain more detailed insights into the fundamental processes of observational learning.

Acknowledgments We would like to thank Lisa Langkamp for her help in conducting the experiment described in this paper.

Funding None.

Compliance with ethical standards

Conflict of interest There was no conflict of interest.

Ethical approval Permission for conducting the research was granted by the university ethics committee. Participants were recruited via the formal channels of the schools. There was informed consent.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Abadie, M., Waroquier, L., & Terrier, P. (2013). Gist memory in the unconscious-thought effect. *Psychological Science*, 24(7), 1253–1259. doi:10.1177/0956797612470958.
- Amadiou, F., Mariné, C., & Laimay, C. (2011). The attention-guiding effect and cognitive load in the comprehension of animations. *Computers in Human Behavior*, 27, 36–40. doi:10.1016/j.chb.2010.05.009.
- Anderson, J. R. (2010). *Cognitive psychology and its implications* (7th ed.). New York: Worth.
- Atkinson, R. K., Derry, S. J., Renkl, A., & Wortham, D. (2000). Learning from examples: Instructional principles from the worked examples research. *Review of Educational Research*, 70(2), 181–214. doi:10.3102/00346543070002181.
- Baddeley, A. D. (2007). *Working memory: Thought and action*. Oxford: Oxford University Press.
- Bandura, A. (1986). *Social foundations of thought and actions: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.
- Bandura, A. (1997). *Self-efficacy. The exercise of control*. New York: Freeman and Company.

- Boltz, M. (1992). Temporal accent structure and the remembering of filmed narratives. *Journal of Experimental Psychology: Human Perception and Performance*, 18(3), 90–105. doi:[10.1037/0096-1523.18.1.90](https://doi.org/10.1037/0096-1523.18.1.90).
- Boucheix, J. M., & Guignard, H. (2005). What animated illustrations conditions can improve technical document comprehension in young students? Format, signaling and control of the presentation. *European Journal of Psychology of Education*, 20(4), 369–388.
- Boucheix, J. M., & Lowe, R. (2010). An eye tracking comparison of external pointing cues and internal continuous cues in learning with complex animations. *Learning and Instruction*, 20, 123–135.
- Brewer, G. A., Marsh, R. L., Meeks, J. T., Clark-Foos, A., & Hicks, J. L. (2010). The effects of free recall testing on subsequent source memory. *Memory*, 18(4), 385–393. doi:[10.1080/09658211003702163](https://doi.org/10.1080/09658211003702163).
- Brucker, B., Scheiter, K., & Gerjets, P. (2014). Learning with dynamic and static visualizations: Realistic details only benefit learners with high visuospatial abilities. *Computers in Human Behavior*, 36, 330–339. doi:[10.1016/j.chb.2014.03.077](https://doi.org/10.1016/j.chb.2014.03.077).
- Butler, A. C. (2010). Repeated testing produces superior transfer of learning relative to repeated studying. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36(5), 1118–1133. doi:[10.1037/a0019902](https://doi.org/10.1037/a0019902).
- Canham, M., & Hegarty, M. (2010). Effects of knowledge and display design on comprehension of complex graphics. *Learning and Instruction*, 20, 155–166. doi:[10.1016/j.learninstruc.2009.02.014](https://doi.org/10.1016/j.learninstruc.2009.02.014).
- Catrambone, R. (1995). Aiding subgoal learning: Effects on transfer. *Journal of Educational Psychology*, 87(1), 5–17.
- Catrambone, R. (1998). The subgoal learning model: Creating better examples so that students can solve novel problems. *Journal of Experimental Psychology: General*, 127, 355–376.
- Chen, C. Y. (2014). Cognitive support for learning computer-based tasks using animated demonstration. *Interactive Learning Environments*,. doi:[10.1080/10494820.2014.936943](https://doi.org/10.1080/10494820.2014.936943).
- Clark, R. E. (2015). Motivating performance: Part 1—Diagnosing and solving motivation problems. *Performance Improvement*, 54(8), 33–43. doi:[10.1002/pfi21503](https://doi.org/10.1002/pfi21503).
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale: Erlbaum.
- de Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2009). Towards a framework for attention cueing in instructional animations: Guidelines for research and design. *Educational Psychology Review*, 21, 113–140. doi:[10.1007/s10648-009-9098-7](https://doi.org/10.1007/s10648-009-9098-7).
- de Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2010). Learning by generating vs. receiving instructional explanations: Two approaches to enhance attention cueing in animations. *Computers & Education*, 55, 681–691. doi:[10.1016/j.compedu.2010.02.027](https://doi.org/10.1016/j.compedu.2010.02.027).
- Dirkx, K. J. H., Kester, L., & Kirschner, P. A. (2014). The testing effect for learning principles and procedures from texts. *The Journal of Educational Research*, 107(5), 357–364. doi:[10.1080/00220671.2013.823370](https://doi.org/10.1080/00220671.2013.823370).
- Eccles, J. S., & Wigfield, A. (2002). Motivation beliefs, values, and goals. *Annual Review of Psychology*, 53, 109–132. doi:[10.1146/annurev.psych.53.100901.135153](https://doi.org/10.1146/annurev.psych.53.100901.135153).
- Ertelt, A. (2007). *On-screen videos as an effective learning tool. The effect of instructional design variants and practice on learning achievements, retention, transfer, and motivation.* (Doctoral dissertation), Albert-Ludwigs Universität Freiburg, Breisgau.
- Giannakos, M. N. (2013). Exploring the video-based learning research: A review of the literature. *British Journal of Educational Technology*, 44(6), E191–E195. doi:[10.1111/bjet.12070](https://doi.org/10.1111/bjet.12070).
- Grossman, R., Salas, E., Pavlas, D., & Rosen, M. A. (2013). Using instructional features to enhance demonstration-based training in management education. *Academy of Management Learning & Education*, 12(2), 219–243. doi:[10.5465/amle.2011.0527](https://doi.org/10.5465/amle.2011.0527).
- Guo, P. J., Kim, J., & Rubin, R. (2014, March 4–5). *How video production affects student engagement: An empirical study of MOOC videos.* Paper presented at the L@S '14, Atlanta, GA.
- Hasler, B. S., Kersten, B., & Sweller, J. (2007). Learner control, cognitive load and instructional animation. *Applied Cognitive Psychology*, 21, 713–729. doi:[10.1002/acp.1345](https://doi.org/10.1002/acp.1345).
- Hassanabadi, H., Robatjazi, E. S., & Savoji, A. P. (2011). Cognitive consequences of segmentation and modality methods in learning from instructional animations. *Procedia-Social and Behavioral Sciences*, 30, 1481–1487. doi:[10.1016/j.sbspro.2011.10.287](https://doi.org/10.1016/j.sbspro.2011.10.287).
- Hodges, N. J., & Coppola, T. (2014). What we think we learn from watching others: The moderating role of ability on perceptions of learning from observation. *Psychological Research*,. doi:[10.1007/s00426-014-0588-y](https://doi.org/10.1007/s00426-014-0588-y).
- Höfler, T. N., & Schwartz, R. N. (2011). Effects of pacing and cognitive style across dynamic and non-dynamic representations. *Computers & Education*, 57, 1716–1726. doi:[10.1016/j.compedu.2011.03.012](https://doi.org/10.1016/j.compedu.2011.03.012).
- Jamet, E. (2014). An eye-tracking study of cueing effects in multimedia learning. *Computers in Human Behavior*, 32, 47–53. doi:[10.1016/j.chb.2013.11.013](https://doi.org/10.1016/j.chb.2013.11.013).

- Jin, S. H. (2013). Visual design guidelines for improving learning from dynamic and interactive digital text. *Computers & Education*, 63, 248–258. doi:[10.1016/j.compedu.2012.12.010](https://doi.org/10.1016/j.compedu.2012.12.010).
- Karpicke, J. D., & Roediger, H. L. (2007). Repeated retrieval during learning is the key to long-term retention. *Journal of Memory and Language*, 57, 151–162. doi:[10.1016/j.jml.2006.09.004](https://doi.org/10.1016/j.jml.2006.09.004).
- Karpicke, J. D., & Roediger, H. L. (2008). *Science*, 319 (The critical importance of retrieval for learning), pp. 966–968.
- Keller, J. M. (2010). *Motivational design for learning and performance. The ARCS-Model approach*. New York: Springer.
- Kester, L., Kirschner, P. A., van Merriënboer, J. J. G., & Baumer, A. (2001). Just-in-time information presentation and the acquisition of complex cognitive skills. *Computers in Human Behavior*, 17, 373–391.
- Kosslyn, S. M., Kievit, R. A., Russell, A. G., & Shephard, J. M. (2012). PowerPoint presentation flaws and failures: A psychological analysis. *Frontiers in Psychology*, 3, 1–22. doi:[10.3389/fpsyg.2012.00230](https://doi.org/10.3389/fpsyg.2012.00230).
- Kriz, S., & Hegarty, M. (2007). Top-down and bottom-up influences on learning from animations. *International Journal of Human Computer Studies*, 65, 911–930. doi:[10.1016/j.ijhcs.2007.06.005](https://doi.org/10.1016/j.ijhcs.2007.06.005).
- Larsen, D. P., Butler, A. C., & Roediger, H. L. (2013). Comparative effects of test-enhanced learning and self-explanation on long-term retention. *Medical Education*, 47, 674–682. doi:[10.1111/medu.12141](https://doi.org/10.1111/medu.12141).
- Lemarié, J., Lorch, R. F., Eyrolle, H., & Virbel, J. (2008). SARA: A text-based and reader-based theory of signaling. *Educational Psychologist*, 43(1), 27–48. doi:[10.1080/00461520701756321](https://doi.org/10.1080/00461520701756321).
- Leppink, J., Paas, F., van Gog, T., van der Vleuten, C. P. M., & van Merriënboer, J. J. G. (2014). Effects of pairs of problems and examples on task performance and different types of cognitive load. *Learning and Instruction*, 30, 32–42. doi:[10.1016/j.learninstruc.2013.12.001](https://doi.org/10.1016/j.learninstruc.2013.12.001).
- Lowe, R. (2004). Interrogation of a dynamic visualization during learning. *Learning and Instruction*, 14(3), 257–273. doi:[10.1016/j.learninstruc.2004.06.003](https://doi.org/10.1016/j.learninstruc.2004.06.003).
- Lowe, R., Schnotz, W., & Rasch, T. (2011). Aligning affordances of graphics with learning task requirements. *Applied Cognitive Psychology*, 25, 452–459. doi:[10.1002/acp.1712](https://doi.org/10.1002/acp.1712).
- Margulieux, L., Guzdial, M., & Catrambone, R. (2012). *Subgoal-labeled instructional material improves performance and transfer in learning to develop mobile applications*. Paper presented at the 9th International Conference on International Computing Education Research (ICER'12), Auckland, New Zealand.
- Mayer, R. E. (2008). *Learning and instruction* (2nd ed.). Upper Saddle River: Prentice Hall.
- Mayer, R. E., & Pilegard, C. (2014). Principles for managing essential processing in multimedia learning: Segmenting, pre-training, and modality principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 316–344). New York: Cambridge University Press.
- McDaniel, M. A., Thomas, R. C., Agarwal, P. K., McDermott, K. B., & Roediger, H. L. (2013). Quizzing in middle-school science: Successful transfer performance on classroom exams. *Applied Cognitive Psychology*, 27, 360–372. doi:[10.1002/acp.2914](https://doi.org/10.1002/acp.2914).
- Merkt, M., Weigand, S., Heier, A., & Schwan, S. (2011). Learning with videos vs. learning with print: The role of interactive features. *Learning and Instruction*, 21, 687–704. doi:[10.1016/j.learninstruc.2011.03.004](https://doi.org/10.1016/j.learninstruc.2011.03.004).
- Moreno, R. (2007). Optimizing learning from animations by minimising cognitive load: Cognitive and affective consequences of signalling and segmentation methods. *Applied Cognitive Psychology*, 21, 765–781.
- Pedra, A., Mayer, R. E., & Albertin, A. L. (2015). Role of interactivity in learning from engineering animations. *Applied Cognitive Psychology*, 29, 614–620. doi:[10.1002/acp.3137](https://doi.org/10.1002/acp.3137).
- Pintrich, P. R., & Schunk, D. H. (2002). *Motivation in education. Theory, research, and applications* (2nd ed.). Upper Saddle River, NJ: Merrill Prentice Hall.
- Plaisant, C., & Shneiderman, B. (2005). *Show me! Guidelines for recorded demonstration*. Paper presented at the 2005 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC'05), Dallas, TX. <http://www.cs.umd.edu/localphp/hcil/tech-reports-search.php?number=2005-02>
- Reisslein, J., Atkinson, R. K., Seeling, P., & Reisslein, M. (2006). Encountering the expertise reversal effect with a computer-based environment on electrical circuit analyses. *Learning and Instruction*, 16, 92–103. doi:[10.1016/j.learninstruc.2006.02.008](https://doi.org/10.1016/j.learninstruc.2006.02.008).
- Reyna, V. F. (2012). A new intuitionism: Meaning, memory, and development in fuzzy-trace-theory. *Judgment and Decision Making*, 7(3), 332–359.
- Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences*, 7(1), 1–75.
- Rheinberg, F., Vollmeyer, R., & Burns, B. D. (2001). FAM: Ein Fragebogen zur Erfassung aktueller Motivation in Lern- und Leistungssituationen. *Diagnostica*, 47(2), 57–66. doi:[10.1026/0012-1924.47.2.57](https://doi.org/10.1026/0012-1924.47.2.57).
- Rieber, L. P. (1990). Using computer animated graphics in science instruction with children. *Journal of Educational Psychology*, 82(1), 135–140.

- Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, 17, 249–255.
- Rosen, M. A., Salas, E., Pavlas, D., Jensen, R., Fu, D., & Lampton, D. (2010). Demonstration-based training: A review of instructional features. *Human Factors*, 52(5), 596–609. doi:[10.1177/0018720810381071](https://doi.org/10.1177/0018720810381071).
- Schittek Janda, M., Botticelli, T. A., Mattheos, N., Nebel, D., Wagner, A., Nattestad, A., et al. (2005). Computer-mediated instructional video: A randomised controlled trial comparing a sequential and a segmented instructional video in surgical hand wash. *European Journal of Dental Education*, 9, 53–58.
- Schoenfeld, A. (1999). Looking toward the 21st century: Challenges of educational theory and practice. *Educational Researcher*, 28(7), 4–14.
- Schwan, S., Garsoffky, B., & Hesse, F. W. (2000). Do film cuts facilitate the perceptual and cognitive organization of activity sequences? *Memory & Cognition*, 28, 214–223.
- Schwan, S., & Riempp, R. (2004). The cognitive benefit of interactive videos: Learning to tie nautical knots. *Learning and Instruction*, 14, 293–305. doi:[10.1016/j.learninstruc.2004.06.005](https://doi.org/10.1016/j.learninstruc.2004.06.005).
- Shippey, S. H., Chen, T. L., Chou, B., Knoepp, L. R., Bowen, C. W., & Handa, V. L. (2011). Teaching subcuticular suturing to medical students: Video versus expert instructor feedback. *Journal of Surgical Education*, 68(5), 397–402.
- Skuballa, I. T., Schwonke, R., & Renkl, A. (2012). Learning from narrated animations with different support procedures: Working memory capacity matters. *Applied Cognitive Psychology*, 26, 840–847. doi:[10.1002/acp.2884](https://doi.org/10.1002/acp.2884).
- Spanjers, I. A. E., van Gog, T., & van Merriënboer, J. J. G. (2010). A theoretical analysis of how segmentation of dynamic visualizations optimizes students' learning. *Educational Psychology Review*, 22, 411–423. doi:[10.1007/s10648-010-9135-6](https://doi.org/10.1007/s10648-010-9135-6).
- Spanjers, I. A. E., van Gog, T., Wouters, P., & van Merriënboer, J. J. G. (2012). Explaining the segmentation effect in learning from animations: The role of pausing and temporal cueing. *Computers & Education*, 59, 274–280. doi:[10.1016/j.compedu.2011.12.024](https://doi.org/10.1016/j.compedu.2011.12.024).
- Spanjers, I. A. E., Wouters, P., van Gog, T., & van Merriënboer, J. J. G. (2011). An expertise reversal effect of segmentation in learning from animated worked-out examples. *Computers in Human Behavior*, 27, 46–52. doi:[10.1016/j.chb.2010.05.011](https://doi.org/10.1016/j.chb.2010.05.011).
- Stiller, K. D., Freitag, A., Zinnbauer, P., & Freitag, C. (2009). How pacing of multimedia instructions can influence modality effects: A case of superiority of visual texts. *Australasian Journal of Educational Technology & Society*, 25(2), 184–203.
- Tabbers, H. K., & De Koeijer, B. (2010). Learner control in animated multimedia instructions. *Instructional Science*, 38, 441–453.
- van der Meij, H. (2014). Developing and testing a video tutorial for software learning. *Technical Communication*, 61(2), 110–122.
- van der Meij, H., & Carroll, J. M. (1998). Principles and heuristics for designing minimalist instruction. In J. M. Carroll (Ed.), *Minimalism beyond the Nurnberg Funnel* (pp. 19–53). Cambridge, MA: MIT Press.
- van der Meij, H., & van der Meij, J. (2014). A comparison of paper-based and video tutorials for software learning. *Computers & Education*, 78, 150–159.
- van der Meij, J., & van der Meij, H. (2015). A test on the design of a video tutorial for software training. *Journal of Computer-Assisted-Learning*, 31(2), 116–132.
- van der Meij, H., & van der Meij, J. (2016). The effects of reviews in video tutorials. *Journal of Computer Assisted Learning*, 32, 332–344. doi:[10.1111/jcal.12136](https://doi.org/10.1111/jcal.12136).
- van Gog, T. (2011). Effects of identical example–problem and problem–example pairs on learning. *Computers & Education*, 57, 1775–1779. doi:[10.1016/j.compedu.2011.03.019](https://doi.org/10.1016/j.compedu.2011.03.019).
- van Gog, T., Kester, L., & Paas, F. (2011). Effects of worked examples, example–problem, and problem–example pairs on novices' learning. *Contemporary Educational Psychology*, 36, 212–218. doi:[10.1016/j.cedpsych.2010.10.004](https://doi.org/10.1016/j.cedpsych.2010.10.004).
- van Merriënboer, J. J. G. (1997). *Training complex cognitive skills: A four-component instructional design model for technical training*. Englewood Cliffs, NJ: Educational Technology Publications.
- Wistia. (2012, May 7). Does length matter? Retrieved from <http://wistia.com/blog/does-length-matter-it-does-for-video-2k12-edition>
- Wouters, P., Paas, F., & van Merriënboer, J. J. G. (2010). Observational learning from animated models: Effects of studying–practicing alternation and illusion of control on transfer. *Instructional Science*, 38, 89–104. doi:[10.1007/s1121-008-9079-0](https://doi.org/10.1007/s1121-008-9079-0).
- Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin*, 127, 3–21. doi:[10.1037/0033-2909.127.1.3](https://doi.org/10.1037/0033-2909.127.1.3).
- Zacks, J. M., & Tversky, B. (2003). Structuring information interfaces for procedural learning. *Journal of Experimental Psychology: Applied*, 9(2), 88–100. doi:[10.1037/1076-898X.9.2.88](https://doi.org/10.1037/1076-898X.9.2.88).